DOCUMENT RESUME

ED 458 065 RC 023 201

AUTHOR Urquiola, Miguel
TITLE Identifying Class Size Effects in Developing Countries: Evidence from Rural Schools in Bolivia.
INSTITUTION World Bank, Washington, DC.
REPORT NO WB-WP-2711
PUB DATE 2000-10-00
NOTE 45p.
AVAILABLE FROM For full text: http://econ.worldbank.org/view.php?topic=10&type=5&id=2597.
PUB TYPE Reports - Research (143)
EDRS PRICE EDRS Price MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Academic Achievement; *Class Size; Elementary Schools; Enrollment; Foreign Countries; Grade 3; Primary Education; *Rural Schools; Small Schools; *Socioeconomic Influences; Socioeconomic Status
IDENTIFIERS *Bolivia

ABSTRACT

Although great interest surrounds class size as a policy instrument, inferences on its effects are controversial. Recent work highlights a particular way to consider the endogeneity issues that affect this variable: class size is often correlated with enrollment, which in turn may be related to socioeconomic status (SES). This paper shows why such correlations may be prevalent in developing countries and that they are indeed a relevant feature in Bolivia. These correlations may partially account for why conventional cross-sectional analyses suggest that larger classes have no effect on or even may raise test scores. The endogeneity issues were addressed through two empirical strategies emerging from institutional arrangements that determine pupil/teacher ratios in rural Bolivia. The first approach exploits a teacher allocation practice that allows schools having classes with more than 30 students to apply for an additional teacher. In these circumstances, enrollment and class size are not "smoothly" related, and their effects may be disentangled. The second strategy considers only remote rural schools with less than 30 students and a single class per grade. In such cases, the enrollment/SES correlations that cause bias are less significant. Analyses focused on grade 3 data. Results of all variations of the research strategies suggest that class size has a negative, significant impact on standardized test scores. (Contains 29 references and 18 statistical figures and data tables.) (SV)

Reproductions supplied by EDRS are the best that can be made from the original document.

ED 458 065

Identifying class size effects in developing countries: Evidence from rural schools in Bolivia

Miguel Urquiola*
Development Research Group
The World Bank

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY
J. Feather
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
2A

*mu26@cornell.edu. For useful comments, I am indebted to David Card, Ken Chay, Paul Gertler, Darren Lubotsky, David Stern, and participants of seminars at Berkeley, Chicago, Cornell, ITAM, Northwestern, The World Bank, and the LACEA meetings. I thank Tito Hoz de Vila, Minister of Education of Bolivia, for providing access to the information used. The institutional support of the Universidad Católica Boliviana, where the empirical work was carried out, is gratefully acknowledged. The findings expressed are the author's, and should not be attributed to the World Bank, its affiliated organizations, or to members of its Board of Executive Directors or the countries they represent.

RC 023201

BEST COPY AVAILABLE

The World Bank Group

Research

go

only in Education

all Research site

go

Title: Identifying Class Size Effects in Developing Countries: Evidence from Rural Schools in Bolivia

Author(s): Miguel Urquiola.

Working Paper No.: 2711

Pub. Date: November 6, 2001

Topics: Education
Rural Development
Social Development

Full Text: Read the full text of this document.

Abstract:

Do smaller classes raise test scores? Evidence from rural schools in Bolivia suggests that they do.

Although class size has attracted great interest as a policy instrument, inferences on its effects are controversial. Recent work highlights a particular way to consider the endogeneity issues that affect this variable: class size is often correlated with enrollment, which may in turn be related to socioeconomic status.

In Bolivia, Urquiola shows, these correlations are significant. Building from institutional arrangements that determine pupil-teacher ratios in rural areas, the author implements two research designs to deal with this issue. The first uses a teacher allocation pattern as an instrumental variable; the second relies on variation from remote schools with a single class per grade. Both suggest that class size has a negative effect on test scores.

This paper—a product of Public Services for Human Development, Development Research Group—is part of a larger effort in the group to explore ways to improve service delivery in education, health, and social protection. The author may be contacted at mu26@cornell.edu.

[More Information]

© 2001 The World Bank Group, All Rights Reserved. Terms and Conditions

BEST COPY AVAILABLE

October, 2000

Abstract

Although great interest surrounds class size as a policy instrument, inferences on its effects are controversial. Recent work highlights a particular way to consider the endogeneity issues that affect this variable: class size is often correlated with enrollment, which may in turn be related to socioeconomic status. This paper shows that these correlations are significant in the case of Bolivia, and building from institutional arrangements that determine pupil/teacher ratios in the rural area, implements two research designs to deal with this issue. The first uses a teacher allocation pattern as an instrumental variable, and the second relies on variation from remote schools with a single class per grade. Both suggest class size negatively affects test scores.
JEL Classification: J0,I2,O1

1 Introduction

Policy makers in developing countries are increasingly interested in raising educational quality, and often turn to economists for guidance on how this might be done. Perhaps surprisingly, however, economic research sheds little light on how educational authorities should allocate their budgets, at least with respect to what inputs they should emphasize. In a survey of developing country studies, for instance, Hanushek (1995) suggests that none of the commonly studied school inputs has a consistently significant impact on test scores.[1] In a review of this research, Kremer (1995) agrees that the evidence on class size, a "popular" input because it can directly and quickly reflect policy, is particularly weak.[2]

The source of disagreements over the effect of class size and other quality measures is simple: these variables are typically correlated with unobserved factors that also influence educational outcomes, a fact that weakens inferences that are not based on credibly exogenous variation. To address this issue, recent research identifies quasi-experimental situations in which the variation in some input is arguably exogenous. This work includes Case and Deaton (1998) in the case of South Africa, and Angrist and Lavy (1999) and Lavy (1999) for Israel.[3] These last two papers highlight a particular way to think about endogeneity in the case of class size.

Namely, this variable is often positively related to enrollment, which may in turn be correlated with socioeconomic status (SES) and achievement. Lavy (1995), for instance, finds an enrollment/SES/achievement link to be an important feature in Israel, and points out that the Coleman (1966) report includes a similar finding for the U.S. Closer to the subject of this paper, Mizala et al. (1999) argue enrollment positively affects test scores in Bolivia, a result they attribute to economies of scale.

This paper first illustrates why such correlations may be prevalent in developing countries, and shows they are indeed a relevant feature in Bolivia. Additionally, it argues they may partially account for why conventional cross sectional analyses suggest that larger classes have no effect on or may even raise test scores, as reviewed in Hanushek (1995). The paper then addresses these endogeneity issues relying on two empirical strategies that emerge from the institutional arrangements that determine pupil/teacher ratios in rural Bolivia.

Using a research design conceptually similar to Angrist and Lavy's (1999), the first approach exploits a teacher allocation practice that allows principals whose schools have classes with more than 30 students to apply for an additional teacher. In the presence of this mechanism, enrollment and class size are not "smoothly" related, making it feasible to identify the latter's effects using the resulting class size function as an instrumental variable. In

[1] These conclusions are broadly similar to those from an even more extensive survey Hanushek (1986) presents for the U.S.

[2] Other surveys are more optimistic about certain inputs' effects. See for instance Fuller (1986) and Fuller and Clarke (1994) for developing countries, and Hedges, Laine and Greenwald (1994) for the U.S.

[3] These efforts are similar to research in the U.S. that explicitly explores the source of variation in some input. See for instance Finn and Achilles (1990), Akerhielm (1995), Boozer and Rouse (1995), Krueger (1999), and Hoxby (1999).

short, this strategy accepts the presence of an enrollment/SES link, but identifies a situation where the effects of class size can be potentially separated from those of enrollment.

The second strategy considers only rural locations so small as to have fewer than 30 students in a single school with only one class per grade. Such institutions are located in remote areas and enjoy a monopoly in educational supply, greatly reducing the scope for parental choice. This, combined with the fact that they do not qualify for additional teachers, implies class size is largely out of their authorities' control. These factors may jointly determine that the enrollment/SES correlations that cause bias in other samples are less significant in this setting. In contrast with the first approach, therefore, this strategy focuses on a situation where it is impossible to isolate class size from enrollment because they are in fact *equal*. Instead, it relies on the fact that in comparisons between remote areas, enrollment/SES correlations may be less important than in broader samples.

All variations of these two approaches suggest class size has a negative and significant impact on test scores. In most cases, the magnitude of the implied effects is not far from those suggested by Finn and Achilles (1990) and Krueger (1999) for the Tennessee STAR experiment,[4] and by Angrist and Lavy (1999) for Israel. Additionally, it is consistent with the early-schooling effects emphasized by Hanushek (1998) and Krueger (1999).

The remainder of the paper is divided into four sections. The next describes the data and provides some background on Bolivia's educational system. Section 3 presents the general empirical framework and descriptive statistics, Section 4 covers the empirical strategies, and Section 5 reviews results. Section 6 concludes.

2 Background and data

Despite a recent transfer of educational infrastructure to more than 300 local governments, Bolivia retains a relatively centralized public educational system, in which the Ministry of Education continues to hire all teachers and set the curriculum. At 100 percent, the primary gross enrollment rate is close to that in countries with comparable income levels. Drop out rates are still relatively high, however, and secondary enrollment ratios are therefore considerably below those of similar countries. Additionally, there is a widespread perception that for any given grade level, a substantial proportion of students is no where near mastery of the grade-appropriate skills.

In response, recent governments have embarked on reforms to improve quality and achieve further gains in access. As part of this effort, the Educational Quality Measurement System (SIMECAL) was introduced in 1993. This system has a testing program as its centerpiece, and is the main source of data for this study.[5] Each year, SIMECAL focuses on one or two grade levels in a sample of schools that includes urban private, urban public,

[4] See Finn and Achilles (1999) and Hanushek (1999) for recent reassessments of the STAR evidence.

[5] See Ministerio de Desarrollo Humano (1997) for an introduction to the testing framework.

and rural public institutions; private schools are very rare in the rural area of Bolivia, and are excluded from this sample.

In addition to gathering language and math test data on each student, the system collects information on his classroom and home background, as obtained from questionnaires administered to: the student himself, his teacher, the head of his household, the school principal, and in the case of rural schools, some community leader. This paper draws mainly on test scores and information collected from students and household heads.

Since its inception, SIMECAL has reached grades 3, 6, 8, and 12. The results presented below focus mainly on the 3rd grade, a choice made for three reasons:

1) Particularly in the rural area, Bolivia still has a schooling supply problem. In a significant proportion of small towns, schools offer only the first three or four elementary grades, so focusing on the lowest grade available increases coverage.
2) The fact that inputs have cumulative effects causes a well known complication in educational production function analyses. Because the 3rd is the lowest of the grade levels available, its test scores are less likely to be "tainted" by previous input availability. This complication also provides an argument for the focus on "small" rural schools that distinguishes one of the identification strategies proposed, since in these cases, between-year changes in enrollment and class size will tend to be small, so that the input of interest is more likely to have been held constant during students' careers.
3) As stated, relatively few students reach secondary education in Bolivia, and differential drop-out rates could bias inferences on class size. This is naturally less of an issue at the 3rd than at the other available grade levels.

3 Empirical framework and descriptive statistics

Historically, developing countries have devoted great efforts to enhance their populations' access to education. Over the years, however, papers such as Behrman and Birdsall (1983) and Case and Yogo (1999) have suggested that the returns to educational investments are not independent of school quality.[6] The emergence of such results, combined with the perception that improving schools can help to eliminate remaining enrollment deficits, explains why researchers have attempted to identify inputs that raise achievement, and why class size, one which policy can quickly influence, has received such attention.

As stated, significant controversy surrounds inferences on these effects, primarily because the extent to which different students enjoy particular inputs is probably related to other characteristics that affect achievement. In the case of class size, recent literature suggests that considering how enrollment/class size and enrollment/SES correlations affect the analysis, is a useful way to think about such sources of bias.

[6] These findings are similar to others that apply to high-income countries, such as those presented by Card and Krueger (1992).

Specifically, if enrollment and class size are positively related, and enrollment is in turn correlated with SES, cross-sectional studies will not reveal the causal effect of class size. Lavy (1995) and Angrist and Lavy (1999) emphasize this point for Israel, and note that an enrollment/SES correlation is also relevant in the U.S., as pointed out by Coleman (1966) and Hedges et al. (1994). They also suggest its presence may bias Akerhielm's (1995) conclusions. In the case of Bolivia, Mizala et al. (1999) use cross sectional evidence to suggest enrollment raises achievement, hypothesizing this may be due to economies of scale.

This section illustrates one way these correlations arise and complicate inferences on class size in the case of Bolivia, and more generally argues they may be relevant in other developing countries. For this purpose, it is useful to note that most empirical work begins with the assumption, often dictated by data availability, that the j^{th} class in school k has size:

$$C_{jk} = \frac{E_k}{n_k}, \tag{1}$$

where E_k is total enrollment in the school, and n_k is the number of classes. For notational simplicity, all these variables are assumed to be *grade specific*.

A first aspect to consider is why enrollment and class size would be related, since, if they are not, any additional enrollment/SES relation will not necessarily pose problems. The setting in which the clearest association arises is among sparsely populated rural locations where schools may have only one 3rd grade, so that enrollment and class size are perfectly correlated. This point is evident in Figure 1, which applies formula (1), plotting class size against enrollment for all schools in the sample. For very small institutions towards the left of the diagram, the two variables are identical. The figure also illustrates that a positive association remains over most ranges of enrollment, if only that which arises from comparing the small and larger establishments.

Beyond the factors that actually cause a positive enrollment/class size link, the key point here is that with this relation present, any additional enrollment/SES correlations will potentially bias inferences on class size. The simplest way to see how these can originate once again involves comparing relatively small and large schools. Because the former will generally be located in rural and the latter in urban areas, any welfare differences between these two realms will give rise to an enrollment/SES correlation. This aspect is relevant in developing countries, where for instance, income per capita is often significantly higher in urban centers than in sparsely populated rural areas.

Figure 2 illustrates this point using information on mothers' educational attainment. Specifically, the figure plots enrollment and the fraction of children in each class whose mothers completed less than five years of schooling, suggesting the following observations:

1) There is a concentration of rural public classes towards the left of the diagram, reflecting the fact that they are more often in schools with low enrollments. In contrast, urban classes, both public and private, are more evenly distributed.

2) A significant proportion of rural classes are in the upper part of the graph, indicating that relatively few of their students' mothers completed more than five years of schooling. In contrast, classes from urban and particularly private institutions are more clustered towards the bottom of the figure, indicating children in these schools come from households with greater educational attainment.
3) To summarize these two points, the figure displays a negative linear relation suggesting that at least as measured by mothers' schooling, children in larger schools come from higher SES backgrounds.

Using 3rd grade class-level data, figures 3 and 4 present an additional illustration of enrollment/SES correlations. The first of these shows relative frequency distributions of enrollment by type of school. As expected, urban classes (the ones covered in the top two panels) display relatively normal distributions, with many schools having between 50 and 150 3rd grade students. More than 70 percent of rural public schools, in contrast, have enrollments below 50.

Figure 4 complements this information showing the distribution of the proportion of students in each class who declare they speak only Spanish. This is a relevant trait in Bolivia, where a significant percentage of the population still uses native languages with varying degrees of intensity.[7] Although the current educational reform stresses multicultural education and a new bilingual testing system is in place, the tests considered here were administered in Spanish, and most instruction still takes place in this language. In general, furthermore, Spanish-only speakers display higher SES than native language users.

This figure shows that a significant majority of children in the urban area, particularly in the private sector, speak only Spanish, but that there is a greater prevalence of indigenous languages in the rural area. For instance more than one in three rural classes contains less than 50 percent Spanish-only speakers, while there are almost no such cases in the urban area. Combined with the information in Figure 3, this data once again suggests a positive enrollment/SES link.[8]

To summarize, figures 2-4 have shown that the enrollment/achievement correlations that Mizala et al. (1999) emphasize may arise not from economies of scale but from enrollment/SES correlations. Figures 5 and 6 suggest they may also reflect positive links between enrollment and the availability of "better" inputs beyond class size. To make this point, they focus on the distribution of the proportion of teachers who declare that more than half of their students have textbooks, and the proportion who are in the top three categories of the pay scale.[9] Combined with the enrollment distributions previously displayed, these data

[7] The two main indigenous languages are Aimara and Quechua. Albó (1995) estimates that slightly more than half of the entire population has some command of one of these.

[8] It might seem surprising that simultaneously, Figure 4 suggests there is a significant proportion of rural classes where Spanish-only speakers are as prevalent as in the urban areas. This heterogeneity in the rural distribution partially reflects geographic variation, namely, the presence of schools from the eastern part of the country, where the prevalence of indigenous languages is low even in agricultural settings.

[9] Bolivia's public educational system has a rigid pay scale which classifies teachers according to seniority

again suggest that children in larger schools generally enjoy "better" inputs.

The preceding examples have all motivated the existence of enrollment/SES correlations as a reflection of rural/urban differences. These links may also arise within urban areas, however, if some schools are more "efficient" in the sense of being better able to raise test scores, and if, additionally, higher SES parents are better able to identify such institutions. This type of "competition" effect, emphasized by Hoxby (1997) in the case of the U.S., may be important in Bolivia because the country has an "open" public educational system, so that admission to a particular school involves no residential requirements.[10] This determines that some public schools perceived to be better by parents, are oversubscribed. To illustrate, in the days surrounding admittance periods, a few schools in major Bolivian cities have some parents physically queuing up to several days to get their children enrolled. To the extent that higher SES or more motivated parents are the ones who engage in such behavior, this may also induce the mentioned enrollment/SES and enrollment/achievement correlations.

These sorting and "competition" effects are probably reinforced by the availability of religious public schools, a trait shared with several Latin American countries. These institutions are generally physically owned and run by a religious denomination (most often the Catholic Church), but their teachers are supplied and paid for by the State. Like all publicly-funded schools, they do not charge direct fees. There is evidence they have more effective incentive systems, that parents perceive them as better, and they are on average oversubscribed more frequently than public schools without religious affiliations.

Regardless of the precise origins of the enrollment/class size and enrollment/SES correlations this section has discussed, it is clear that in their presence it may be impossible to identify class size effects using cross-sectional variation. In particular, these links may bias estimates upward, suggesting this variable may have less of a negative effect than might actually be the case. The descriptive statistics in Table 1 illustrate this point. This table presents information on the whole sample of schools, and also on each type of institution: urban private, urban public, and rural public.

A first point is the ordering of average test scores in both subjects: students in urban private schools place highest, followed by those in urban public and rural public institutions. As expected, the table also shows that urban and particularly private school students display "better" socioeconomic characteristics and enjoy "higher quality" inputs. The key point is that at the same time, average class size is highest in the urban private sector, followed by the urban public and rural public cases, *so that a comparison of means would suggest larger classes in fact produce better outcomes.*

In this context, the fact that higher SES children are often in larger classes may seem surprising. It is important to realize, however, that the class size levels children experience are the result of a complex decision process in which parents may consider not only this

and formal qualifications. See Urquiola et al. (2000) for detail.

[10] There have recently been attempts to introduce residential controls, but anecdotal evidence indicates these are either not enforced or ineffective.

variable's level (or even that of other educational inputs), but also aspects such as their children's peer group composition. If one adds to this the fact that schools have cost-related incentives to keep class sizes high, it is clear that no straightforward predictions on the relation between, say, income, and class size will be available.

To summarize, the simple information presented in this section shows that biases arising from a combination of enrollment/class size and enrollment/SES correlations will typically affect inferences about class size in countries like Bolivia. The following section introduces the empirical strategies used to deal with this problem, and to address the endogeneity of class size more generally.

4 Empirical strategies

Most work on class size begins from a reduced form specification which assumes that the test score T of an individual i in class j, school k, and town l is given by

$$T_{ijkl} = \alpha + \beta_C C_{jkl} + \hat{X}_{ijkl}\beta_{\hat{X}} + \check{X}_{jkl}\beta_{\check{X}} + \tilde{X}_{kl}\beta_{\tilde{X}} + \bar{X}_l\beta_{\bar{X}} + \epsilon_{ijkl}, \tag{2}$$

where $\hat{X}_{ijkl}$ is a vector of students' socioeconomic characteristics, $\check{X}_{jkl}$ denotes traits specific to given classes, such as teachers' qualifications, $\tilde{X}_{kl}$ refers to school characteristics, and $\bar{X}_l$ stands for town-level variables, like population size (C_{jkl} is as defined in formula 1, page 6). The previous section illustrated that in any practical application, the X vectors are likely to exclude key characteristics that affect achievement, and being correlated with class size, will bias estimates of β_C.

Another way of stating this is that while class size may be given simply by the ratio of enrollment and the number of classes offered, as in (1), in reality both of these are affected by socioeconomic characteristics:

$$C_{kl} = \frac{E_{kl}(\hat{X}_{ijkl}, \tilde{X}_{kl}, \bar{X}_l, \epsilon_{ijkl})}{n_{kl}(E_{kl}, \hat{X}_{ijkl}, \tilde{X}_{kl}, \bar{X}_l, \epsilon_{ijkl})}. \tag{3}$$

Placing these variables in the numerator of this expression makes explicit the possible enrollment/SES correlations illustrated in the previous section. That they also belong in the denominator is evident in Figure 1 which shows that not only does class size vary with enrollment, but schools with identical enrollment levels often choose to offer different numbers of classes, a decision which may naturally reflect traits like income levels.

Both (2) and (3) make clear that reliably identifying class size effects will necessitate experimental or quasi-experimental evidence. The previous section suggested that a way of thinking about this requirement is that one must have situations where either: i) class size and enrollment are unrelated, or at a minimum, ii) enrollment and SES correlations

are less prevalent than in the aggregate sample.[11] This section argues that the institutional characteristics of educational provision in the rural area of Bolivia make these approaches feasible, and presents two empirical strategies that implement them, one at a time.

Before proceeding to these identification issues, however, it is relevant to mention a measurement problem. Boozer and Rouse (1995) emphasize that calculating class size by equation (1), as is actually done in this study and is illustrated in Figure 1, can result in bias. They illustrate this possibility by noting that some school districts in the U.S. have compensatory policies that place worse-performing students in smaller classes, so that using actual class sizes rather than some school average is necessary for accurate inferences.

This measurement problem is fortunately not a major issue in the present setting. First, the SIMECAL system provides both measure (1), collected from school officials, and the actual number of pupils teachers declare are in their classes. Figure 7 shows that these two measures are in fact very similar: most observations are very close to a hypothetical 45 degree line that would indicate they are identical. This is particularly the case among the rural schools that are the focus of this study.

This reflects that in Bolivia, schools often have only the number of teachers they strictly need to cover the grade levels offered. In the extreme, schools in smaller communities will have only one class per grade level, and will rarely have a principal or any administrative staff.[12] As in many low-income countries, therefore, making subtle adjustments to actual class sizes, or offering compensatory education, will rarely if ever be in public schools' choice set. Additionally, the focus here is on primary education, where class size tends to be a more clearly defined concept than in later schooling stages.[13]

If based on this logic one ignores the measurement problem, then C need not have a j or class subscript, and (2) takes the form:

$$T_{ijkl} = \alpha + \beta_C C_{kl} + \hat{X}_{ijkl}\beta_{\hat{X}} + \check{X}_{jkl}\beta_{\check{X}} + \tilde{X}_{kl}\beta_{\tilde{X}} + \bar{X}_l\beta_{\bar{X}} + \epsilon_{ijkl}, \tag{4}$$

where class size, conditional on grade level, will depend only on the school students are in. This will certainly be the case in small institutions which have one class per grade. With this measurement issue aside, the following sections present the mentioned empirical strategies.

4.1 Teacher allocation patterns as an instrumental variable

Following Angrist and Lavy (1999), a first approach exploits the fact that teachers are allocated in a centralized manner in Bolivia. Under prevailing practice, when a given district

[11] SES and class size could of course still be correlated through mechanisms other than those operating through enrollment.

[12] There will generally be a "district principal" who serves several small schools from a larger town. As stated, the system is still relatively centralized, and these districts have little autonomy.

[13] The data consistency displayed in Figure 7 also matters to the extent one may be concerned that administrative data in developing countries is not reliable. Anecdotal evidence indicates that educational reform efforts have resulted in more accurate reporting in Bolivia.

can demonstrate one of its schools has more than 30 students per teacher in a given grade, it can apply to the Ministry of Education for another teacher, thus lowering the class size at that school (for the given grade) significantly. For now, the discussion assumes a situation in which this rule binds without exception, that is, districts always request a teacher, and one is always granted, when they reach a class size of 30. Further below, this section presents evidence on the extent to which this procedure actually binds.

Mathematically, this practice determines that the number of classes a public school can offer, n^*_{kl}, is a discontinuous function of its enrollment:

$$n^*_{kl} = f(E_{kl}) = \frac{E_{kl}}{int(\frac{E_{kl}-1}{30}) + 1}. \tag{5}$$

Graphically, Figure 8 illustrates the impact of this regulation on the relation between enrollment and class size. Up to 30 students, these two variables take on identical values; thereafter, (5) causes class size to fall discretely at enrollment intervals of 30 students.

In terms of an identification strategy, the relevant consequence is that if the teacher allocation procedure is always binding, equation (3) becomes

$$C_{kl} = \frac{E_{kl}(\hat{X}_{ijkl}, \tilde{X}_{kl}, \bar{X}_l, \epsilon_{ijkl})}{n^*_{kl}}, \tag{6}$$

the key change being that neither X nor ϵ now affect schools' choice of n. Graphically, this is seen comparing figures 1 and 8. When a rule like (5) applies, two schools with the same enrollment level will never have different class sizes.

This change does not succeeded, however, in eliminating X and ϵ from the class size determination equation, since they still act through E_{kl}. Nevertheless, *as long as achievement and enrollment are "smooth" functions of these two variables*, (5) still provides useful variation. The argument behind this, as put forth by Angrist and Lavy (1999), is an application of the quasi-experimental regression discontinuity design introduced by Campbell (1969), who analyzed how to identify the effect of a treatment variable (in this case, class size) assigned on the basis of an observed continuous measure (in this case, enrollment), called the selection variable. Because the selection variable itself may be related to the outcome (e.g. through the enrollment/SES correlations described above), the treatment variable could be statistically related to the outcome even if the relationship is not causal.

The logic is that if enrollment is smoothly related to SES, observations in narrow intervals around discontinuities provide a quasi-experimental setting because the sample right above each "tooth" in Figure 8 should be very similar to the one right below. It is therefore plausible that any variation in their outcomes is caused by differences in the treatment variable. In terms of equation (4), using this rule involves utilizing (5) to instrument C_{kl}, running a regression like:

$$T_{ijkl} = \alpha + \beta_C \hat{C}_{kl} + X_{ijkl}\beta_X + \beta_{E^*} E^*_{kl} + \epsilon_{ijkl}, \tag{7}$$

where $\hat{C}_{kl}$ denotes instrumented class size, E^*_{kl} is some "smooth" function of enrollment, possibly $ln(E_{kl})$ or enrollment itself, and the control variables have been combined into X_{ijkl} for simplicity.

To check how useful the mentioned teacher allocation procedure can be as an instrumental variable, a simple first step is to observe whether average test scores, when plotted against enrollment, display an up and down "saw tooth" behavior *inverse* to that which characterizes the class size function in Figure 8. In the case of Israel, for instance, Angrist and Lavy (1999) show that this is clearly the case, pointing out it is difficult to attribute such behavior to anything but the variation in class size induced by the class size rule they use in that context.

For the present data, this exercise is shown for the rural area in Figure 9. For reasons discussed later in this section, the teacher allocation procedure described is not binding in urban areas. The top portion of this illustration graphs average actual and predicted class sizes by enrollment levels, and the bottom plots averages of math and language scores, also by enrollment levels. The figure suggests several observations.

First, the actual does not always track the predicted class size. Rather than starting to assign new teachers at grade-specific enrollment levels of 30, the data suggest authorities seem to wait until about 38 students, which may reflect resource constraints. This also illustrates that the teacher allocation mechanism reflects practice rather than a legally binding rule. There are aspects that reflect a commitment to implicitly limiting class sizes, however. For instance, the Social Investment Fund, the agency in charge of building schools in the rural area, explicitly designs its classrooms for 30 students. This guideline, mandated by the Ministry of Education, is also observed by the main Non-Governmental Organizations active in school construction.

What is important for the purposes of this paper, is that at a *given* level (about 38 students) actual class sizes seem to drop for most schools, i.e., by enrollments of 40 most schools have been assigned a new teacher, creating the clear discontinuity observed in this range. The absence of a discontinuity would be evidence that teacher allocation is more endogenous than the straightforward application of a teacher allocation rule would suggest. The fact that not all new teachers are assigned at exactly the same enrollment level, nevertheless, makes it impossible to completely discount biases. Additionally, some further "spikes" observed at greater enrollment levels may also be a cause for concern.

There is another perhaps clearer discontinuity at about 65 students, but beyond enrollments of about 72 students the teacher allocation rule ceases to predict class size effectively. The greater variability observed in part reflects smaller sample sizes in this data, but may also be due to the fact that towns with such enrollment levels begin to become more "urban". Considering the rural area as a whole, nevertheless, the described teacher allocation procedure is a relevant practice which accounts for 40 to 50 percent of the variation in class sizes, an explanatory power similar to that observed in the Israeli case.

Turning to the test scores, even though these are *not* smoothed and the samples are

small relative to those used in other studies, the figure displays a suggestive pattern generally inverse to that observed for class size. In the first segment, that which covers enrollment levels from zero to about 38, there is a clear negative relation between test scores and class size. Much of this arises from the small rural schools that are the focus of the next identification strategy, so a more complete discussion of this evidence is postponed.

More importantly for the current research design, at almost the precise level where the teacher allocation rule causes class sizes to fall discretely, test scores jump abruptly. Although the behavior of scores between the two discontinuities is more erratic, the overall pattern is still suggestive of an inverse relation. Additionally, test scores jump again at about the enrollment level (66-68 students) where there is another discrete reduction in class sizes. Beyond this last enrollment level, test scores begin to decline again with enrollment, although not as clearly as in the previous segments. This may partially reflect enrollment/SES correlations that begin to become more important as towns become larger and more urban, an aspect further discussed below. A final point is that the behavior of both tests is rather similar, which is not always a feature of studies that cover more than one subject score.

As mentioned, this identification strategy focuses only on the rural area, Although excluding the urban realm eliminates a significant part of the sample, this restriction is necessary because the practices described do not bind in this case. This much is already evident in the descriptive statistics presented in Table 1, which show that the *mean* class size in urban public schools is close to 36 (with a standard deviation of 6), a level that is inconsistent with a similar rule's application. The mean class size in the rural area, in contrast, is 23 students, with a standard deviation of 8.

The are no immediately apparent reasons for this asymmetry in the allocation of teachers, or for the between-area differences in per pupil expenditures it implies.[14] Nevertheless, the following points are relevant to an explanation:

1) The existence of sparsely populated rural areas implies that educational authorities must necessarily be willing to tolerate lower class sizes in this segment in order to achieve given target national enrollment ratios. While this helps explain the urban-rural differences in average class sizes, it does not account for why the teacher allocation guidelines would be differentially enforced.
2) As indicated, Bolivia has an "open" public school system, so that admission to a particular school involves no residential requirements. This makes urban enrollments hard to predict and may render any rule hard to apply, even if that is what the educational authorities wanted to do.
3) The fact that the *mean* urban class size is significantly larger than 30, however, calls into question whether authorities actually pursue this goal, which may in turn have to do with resource constraints that have different implications in each area.
4) Specifically, the *kinds* and *costs* of teachers the Ministry assigns to each area are dif-

[14] As in many developing countries, the majority of Bolivia's education budget is devoted to salaries, so that student/teacher ratios are an important part of the overall resource allocation picture.

ferent, which may explain its apparent greater willingness to spend on teachers and implicitly enforce an allocation rule in the rural area. Differences in the types of teachers are relevant because there is a shortage of certified instructors. Because this situation is most acute in rural regions, regulations allow authorities to appoint "interim" teachers in these areas. These are individuals who do not have a teaching certificate, and will often not have finished high school. In many cases, they are natives of the community they serve. All these traits determine they are also at the bottom of the national pay scale, and are therefore more available and less expensive than the certified instructors authorities more often assign to the urban area.[15]

Combined, these factors may explain why teacher allocation seems to be more systematic in the rural area, and why the average class size is lower in this realm.

In sum, this figure suggests than an instrumental variables approach like that suggested by (7) may produce results different from those observed in conventional OLS specifications. These are presented in Section 5 below, where the the allocation rule used is that which effectively binds, with a discontinuity at about 38 students.

4.2 Focusing on small towns

As Angrist and Lavy (1999) point out and Hoxby (1999) emphasizes, the IV approach suggested by (7) may be appropriate only at the *discontinuities* induced by a given class size rule. Regarding the present data, this implies the strategy may only be valid when considering observations within some narrow interval of the enrollment levels that cause abrupt drops in class size, since it is only around such discontinuities that any "smooth" enrollment/class size link is really severed. Actually implementing this adjustment, however, is likely to cause the loss of many if not most observations in cross-sectional data sets.

Nevertheless, to the extent that the research focus is on rural schools, it may not make sense to consider only the variation within such bands. Specifically, there may be valuable information in the observations up to the first "tooth" of the enrollment/class size relation, *when that arises from towns so small as to have a single public school with only one class per grade*. Such communities will be referred to as "small towns" in the subsequent discussion, and their value in this setting arises from a number of observations:

1) Given the teacher allocation procedure and prevalent educational supply restrictions, public schools in these locations have a monopoly on formal schooling, and parents essentially have no choice between institutions or classes. This reflects that private schools are very rare in the rural area and, when they exist, are almost always located in large towns that can support a religious institution alongside one or more public ones.

[15] There are also "interim" instructors in cities, but their prevalence is much lower. This reflects the fact that most teachers seem to prefer to live in cities, and that those that have degrees have greater choice on their geographical assignment.

2) The statement that parents cannot exercise choice must be qualified because they could always move in reaction to an observed class size. While this is possibly an important concern for Hoxby (1999) given the mobility and proximity that exists between U.S. school districts (particularly in a small state like Connecticut), and for Lavy (1999), given the population density and relatively small area of Israel, it may be less of an issue in rural Bolivia. In this case, migrating does not imply switching areas of residence while possibly keeping the same job, but may require attempting to sell land on which there are often no clear property rights, and moving to a relatively distant location, an aspect that is reinforced because schools are generally few and far apart. This motivation is similar to that used by Case and Deaton (1998), who exploit Apartheid-imposed mobility restrictions to identify resource effects in South Africa.
3) Because schools in small towns clearly do not qualify for additional teachers, school officials will have almost no latitude to alter class sizes or "campaign" for more instructors. This may be an important advantage relative to the previous strategy, particularly if one is concerned about possible manipulations of the teacher allocation rule around the discontinuity-inducing enrollment levels.

In light of these observations, the class size a student in a small town experiences will be almost solely determined by his or her cohort size. In terms of Angrist and Lavy (1999), *this is a situation where the assignment and treatment variables are one and the same, but may nevertheless have some relevant exogeneity characteristics.* Formally, this approach entails running an OLS regression of the form:

$$T_{ijkl} = \alpha + \beta_C C_{kl} + X_{ijkl}\beta_X + \beta_{E^*} E^*_{kl} + \epsilon_{ijkl}, \quad where \quad 0 < E_{kl} < 31. \tag{8}$$

where the focus on small rural locations leaves little room for biases arising from within or between-school differences in a given community. Following what would be a strict application of this rule, this specification calls for placing the cutoff at 30 students, while Figure 9 could be interpreted as arguing for a higher cutoff, perhaps around 38 students. The results below use the lower benchmark, but were not found to be sensitive to this choice; if anything, the key coefficients are more significant with a higher cutoff point. The lower cutoff was used, however, because as discussed further it makes it more likely that small towns (with only one class per school and one school per town) are being selected.

This use of between-town variation, of course, does not eliminate all possible sources of bias. Even in this small town sample, larger locations will tend to have larger cohorts, enrollments, and class sizes. Additionally, these could systematically attract better or worse teachers than smaller ones. To express this point formally, note that under the above arguments, class size will be given not by (3) but rather by:

$$C_l = \frac{E_l(\bar{X}_l, \epsilon_l)}{1}. \tag{9}$$

where n is now set to 1. The absence of subscripts j and k reflects there is only one

class and one school in each town indexed by l. Class level characteristics, $\tilde{X}_{kl}$, have been eliminated for the same reason, whereas the absence of student and parental traits, $\hat{X}_{ijkl}$, reflects the assumption that parents can no longer influence achievement through their choice of school. The presence of $\bar{X}_l$, however, indicates that inter-town heterogeneity could still bias inferences if it is correlated with factors that affect achievement. This is another way of stating that there could still be enrollment/SES/achievement correlations, even if the enrollment variation arises from differences in towns' sizes rather than differences between schools. While this is certainly possible, the results below suggest that the biases that arise from this are probably less severe.

4.3 Combining the allocation rule and the focus on small towns

If the above arguments hold and the variation among small towns is meaningful, a logical extension is to combine the two previous research designs. In this case, the strategy focuses only on rural schools with enrollments less than or equal to 30 (those in small towns), on the one hand, and on those with enrollments within narrow bands of levels that trigger discontinuities in the enrollment/class size relation, on the other. This reflects that having small town schools as a base provides a sample size large enough to then restrict the remaining observations to those which provide "higher quality" variation. Specifically, the approach in this case is,

$$T_{ijkl} = \alpha + \beta_C \hat{C}_{kl} + X_{ijkl}\beta_X + \beta_{E^*} E^*_{kl} + \epsilon_{ijkl}, \quad where \quad 0 < E_{kl} < 45 \quad or \quad (10)$$

$$63 < E_{kl} < 68$$

and the instrumental variables technique binds only for schools with enrollments greater than 30. In practice, the results were not found to be significantly sensitive to the exact placement of these cutoff points.

The next section presents results on these identification strategies. It also features brief evidence on a couple of additional methods that address sources of endogeneity the two approaches discussed here do not. Because these are not central to the paper, the methodological details are briefly discussed within the results section itself.

5 Results

For comparison purposes, this section first presents results that apply to all schools considered, including those in the urban samples. The regressions then implement the identification strategies described in the previous section, which, as stated, focus only on rural area schools.

5.1 General results

To introduce results on class size and enrollment, Table 2 presents simple cross sectional specifications for the full sample and each type of school: urban private, urban public, and rural public. Panel A has class size as the key independent variable. In this case, column 1 suggests larger classes in fact produce *higher* test scores, a statistically significant result which is not surprising given the correlations discussed above. Such a positive point estimate is also observed among urban private schools in column 3. When column 2 focuses on the full sample and adds dummies for whether schools are in the private sector or the rural area, the coefficient on class size becomes negative and significant. The result is even greater in magnitude when the regression considers only public schools in columns 4 and 5. The signs of all coefficients are the same in panel B, except that the focus is on enrollment as the key independent variable. As expected, the results suggest enrollment positively affects achievement in the aggregate sample.

Before proceeding to the rural-only results, Table 3 makes the point that results in all areas behave in a manner consistent with Hanushek (1995). This table once again considers the full sample, urban private, urban public and rural public schools. For each of these groups, the first regression simply replicates the corresponding simple specification observed in Table 2. The second one adds a number of socioeconomic controls, which have the effect of lowering the coefficients on class size and rendering them insignificant at the 5 percent level.[16]

The socioeconomic control variables, in contrast, almost always have effects in the expected direction and are highly significant in the aggregate sample. Students who are native Spanish speakers have higher scores; working tends to lower performance, and mothers' schooling to improve it. Two characteristics, whether households have access to sewerage and phone connections, turn out to be highly significant, perhaps because they proxy for income and urban status.

5.2 Rural area results: the teacher allocation rule

With these introductory results as a benchmark, the remainder of the regressions focus on different subsets of rural schools only. Additionally, while the previous results covered only language test scores, the subsequent cover math as well. Introducing the use of the teacher allocation rule and following Angrist and Lavy (1999), Table 4 presents reduced form equations which use the *predicted* class size, computed using the class size function, as the key independent variable, where this was assumed to entail the introduction of another teacher when enrollment reaches 38 students. While this precise level is arbitrary and point estimates display some variation, none of the conclusions listed below are sensitive to placing

[16] Although the analysis here centers on class size, using the same data Urquiola et al. (2000) find that in cross-sectional settings, most conventional educational inputs like teacher's experience and textbook availability, cannot be shown to have any systematic effects on test scores either.

the critical level at 37 or 39.

Column 1 contains a first stage regression, where the dependent variable is the actual class size. Not surprisingly, the coefficient of interest is highly significant, and even in this simple specification, R^2 is above 0.5. If class-level observations are used instead, R^2 is about 0.44, a level comparable to that found for Israel by Angrist and Lavy (1999), who also use data at this level of aggregation. Using class level observations produces class size effects which are greater in magnitude, for reasons discussed below. To take advantage of the more detailed data, however, almost all the results presented here are at the individual level, with standard errors adjusted for clustering at the class level.

Columns 3-6 present analogous specifications, except that the dependent variables are now language (columns 3 and 4) and math (columns 5 and 6) scores. The coefficients are significant in all cases, and their magnitude is always greater than that observed in the "simple" OLS regressions in Table 2. Importantly, adding socioeconomic controls does not alter the coefficients on class size as much as was the case before, and does not render them insignificant. An exception is the final specification for math scores, where the estimate is only significant at the 10 percent level. In contrast, the socioeconomic controls are not as important in this setting, which in part may reflect smaller variances for these variables within the rural area.

Table 5 presents the actual instrumental variables results, based on the first stage regression presented in Table 4. For comparison, columns 1-2 (language) and 5-6 (math) present simple OLS specifications. For each test score, the table features two IV regressions, one without control variables, and one that includes enrollment and socioeconomic characteristics. For both tests, and in the simplest univariate specifications, the coefficients on class size are negative and significant in both the OLS and IV regressions. When control variables are added, however, the significance in the OLS case is lost, whereas the IV coefficients not only remain significant but become greater in magnitude, suggesting larger classes result in lower achievement. Once again, for the case of math this is only true for the 10 percent level. Similar results emerge when other specifications of enrollment, such as its log, are used among the control variables.

As stated above, the use of the teacher allocation rule is based on the argument that it "separates" the effects of enrollment and other socioeconomic characteristics from those of class size, making it feasible to estimate the causal effect of the latter. If this strategy indeed identifies exogenous variation, these results confirm the assertion that simultaneous enrollment/class size and enrollment/SES correlations cause bias in the usual cross-sectional estimations, hiding a negative and significant class size effect.

5.3 Rural schools: focusing on small schools

In an earlier section, this paper argued that concentrating on "small" towns, those with a single school and 3rd grade enrollment below 30, might identify useful variation in class

size. In these cases, short of making what might be a rather costly move, parents cannot react to a class size level that arises from the size of their child's birth cohort, and schools are not large enough to be candidates for a supplementary teacher.

Table 6 presents results for this sample. Because this strategy does not rely on an instrumental variables technique, the table covers OLS results only. For each test, the first two columns present estimations on the whole rural sample as a benchmark, and the following two focus only on schools with enrollments less than or equal to 30. The coefficients on class size are uniformly significant in the "small town" sample, and in contrast to all other samples in OLS specifications, their magnitude is barely affected by the introduction of socioeconomic controls.

In short, the key finding here is that class size has a negative and significant effect among rural schools in small towns, and that this may be because enrollment/SES correlations are less important in this context. This may not be surprising to the extent that one would expect enrollment/achievement links to be less significant when they arise from differences in towns' population sizes rather than from parental choice. Additionally, of course, the magnitude of small towns' enrollment cohorts will not be determined just by their size, but also by a random element. This point is explored in greater detail in a subsequent section.

The sample considered in making this finding is, of course, different from the one used earlier, and even if the identification strategy does isolate useful variation in class size, these effects might be biased up or down relative to those that would prevail if such variation were identified for, say, a random sample of the urban population.

5.4 Rural schools: combining small schools with discontinuities

A natural extension is to combine the two sources of variation used so far: that arising from the focus on "small" schools, and that related to the use of a class size rule. Because the former sample provides an adequate base from which to build, the latter can be restricted to only those schools within narrow intervals associated with discontinuities in the enrollment/class size relationship. In a sense, this adds only the "higher" quality variation from the IV strategy to the data from the "small" schools approach.

To this end, Table 7 implements regression (10) for language and math scores. For comparison, in each case the first two columns present OLS regressions for this sample, with the remaining two focusing on IV specifications. In the latter, the coefficients on class size become more negative and significant than those in Table 6, which applied to small schools only, or Table 5, which presented the IV results. One point these comparisons suggest is that the significant effects found among the small schools, a very particular set of institutions, might not necessarily be overestimates of those one would observe in larger institutions.

5.5 Possible sources of bias and further results

The results presented thus far consistently suggest that once strategies to address the endogeneity of class size are put into place, this variable is found to have negative effects on achievement. These approaches obviously do not substitute for experimental evidence, and will therefore still leave open routes for bias. This section discusses some of them and, where possible, presents some empirical exercises that provide information as to their importance.

5.5.1 Geographical influence areas and school quality

Up to this point, the discussion has assumed that rural schools are located in and draw students from towns. Among rural schools, however, the realm from which enrollment originates will often encompass a larger, sparsely populated agricultural area from which children walk to school each day. One reason this matters is that as the distance to school becomes greater, there may be households who deem their children's trip worthwhile only if the education offered is at least of a given quality. Additionally, even though schools tend to be far apart, one cannot rule out the existence of children who live at the margin of two schools' areas of influence, whose parents could therefore send them to the one they perceived to be better. This implies that even in rural areas with theoretically little scope for choice, schools' enrollment and class sizes, therefore, may actually be a function of their performance on standardized tests.

While this is a relevant issue, it is important to note that such an effect would tend to bias the estimated class size effects upward, making negative and significant results like those observed here harder to encounter. An indirect test of this issue involves augmenting the above regressions with some variable that might capture quality and "school effects" more generally. One possibility is to take into account the fact that schools that have a 6^{th} in addition to a 3^{rd} grade might be of higher quality. This could be due to the fact that the institution may be older and have more experience, or that the district principal or local teachers have been more effective in "lobbying" for the provision of additional instructors and infrastructure. This idea was implemented by adding a dummy for 6^{th} grade availability to some of the above regressions. While the coefficient on this variable is sometimes significant, it does not have a consistent sign and barely affects the key class size estimates.

5.5.2 Inter-town differences

Another way to deal with this problem is to focus only on the part of enrollment variation that is due to differences in cohort sizes *between years, within towns.* Despite the fact that this significantly restricts the variation in class sizes, it may allow further controls for any enrollment/SES correlations that remain in comparisons between small towns. For instance, it would seem unlikely that teachers with given characteristics would not want to live in small towns to teach 3^{rd} grade, but would not mind teaching 6^{th} grade in them. Further, to the extent that Bolivia has relatively high fertility rates, in many cases the

students in 3rd and 6th grades will come from the same households, though the data does not allow one to identify these links.

This approach is similar to that used by Hoxby (1999). The problem it poses in the present context is that repeated observations are not available for the same grade in the same town. Analogous variation, however, can still be obtained if one focuses on cross-cohort differences between the 3rd and 6th grade, both grade levels covered in the data.[17]

The main disadvantage of this approach comes from the possible presence of differential drop-out rates that can bias results in unpredictable ways. Towns with better schools, for instance, will be those in which children tend to remain in school and reach the sixth grade. All else equal, such schools will probably have higher scores than those in which students tend to drop out, but they will at the same time have larger class sizes. On the other hand and in comparison to "high drop out" areas, they will also be retaining a larger proportion of poorly performing students.

Implementing this strategy requires selecting only "small" towns that contain schools that offer both 3rd and 6th grade instruction, which unfortunately further restricts the sample. Additionally, both the 3rd and the 6th grade must have less than 30 students. Table 8 presents the results for this approach, focusing only on language scores and using observations at the class level. Columns 1-3 present regressions without socioeconomic controls, while columns 4-6 incorporate these. Within each group, the first regression contains only class size as the key independent variable. In order to capture constant differences between the two grades' average scores, the second adds a dummy indicating whether the observation is at the 6th grade level. The coefficient on this dummy is uniformly positive but not significant. Finally, the third column within each group adds town dummies, with which the results rely only on within town, cross-cohort variation in class sizes. The final specifications do suggest negative and significant effects.

5.6 Interpretation

Perhaps due to the controversy that surrounds this input, most work on class size focuses on the direction and significance of its effects. From a policy perspective, of course, the *magnitude* of any potential gains from class size reduction is what matters. To address this issue, some papers use a measure called the *effect size*: the implied change in test scores that results from a given change in class size, divided by the standard deviation of the score. Using a one standard deviation change in class size, the above tables have presented this statistic in brackets. Finn and Achilles (1990) point out that the effect size will not be independent of whether individual or class average observations are used for test scores. Because the variance in the former is naturally greater, studies which use individual level observations will tend to find smaller effects.

[17] Forthcoming data will include also include 1st grade results. To the extent that these can be linked to the schools from which these 3rd and 6th grade results were drawn, this will allow a further check on these results.

Table 9 summarizes the effects found in this paper, comparing them with some encountered in previous research.[18] The first two rows show the effect sizes found by Finn and Achilles (1990) for the STAR experiment, and Angrist and Lavy (1999) for the case of Israel. Both of these refer to the proportion of a standard deviation change in test scores that would be caused by an *eight pupil* reduction in class size. Both results display the same orders of magnitude, although the Israel results are towards the lower limit of the STAR range.[19]

The remaining rows present the effect sizes found in this study for the two identification strategies, and for their combination. This study has used "one standard deviation" effect sizes which are not strictly comparable to the "eight student reduction" ones presented for Finn and Achilles (1990) and Angrist and Lavy (1999). Nevertheless, because one standard deviation corresponds to between eight and nine students in the rural data (see Table 1), the interpretation is similar. With this caveat, the table shows that the IV results produced here are slightly larger than those found by Finn and Achilles (1990) when the student distribution is the benchmark, but within their range when the class-level standard deviations are used instead. The small school estimates are not as large, and are therefore closer to those found for Israel.[20]

This table gives a general notion of how the estimates presented compare with others. A more intuitive interpretation is possible by comparing the suggested class size impact with that which parental education implies. To illustrate, suppose class sizes in rural Bolivia were reduced by half a standard deviation, from about 23 to 19 students. The impact on test scores Table 5 suggests is roughly similar to that suggested by the dummy for whether the mother is a high school graduate.

This is a significant impact, an aspect consistent with Moesteller (1995), who emphasizes that even if effect sizes seem small, they can be equivalent to substantial movements along the test score distribution. For instance, a 0.30σ effect size in the case of rural schools would mean that an eight student reduction in class size would be sufficient to move a student from the 25^{th} percentile to the median of the language test score distribution for the entire sample.

A further consideration is that especially at the grade level considered here, higher test scores may reflect the acquisition of skills less "marginal" than those measured in developed countries. Specifically, adequate performance in the language test at the 3rd grade level may reflect a Bolivian rural student is about to acquire the literacy skills requisite for effective citizenship in a Spanish-speaking society. This aspect also suggests some caution in reading the comparisons presented in Table 9.

[18] This table is partially based on Angrist and Lavy's (1999) discussion of these topics.

[19] Krueger (1999) presents generally similar results for the STAR data.

[20] Akerhielm (1995) and Boozer and Rouse (1995) present higher grade results based on NELS for the U.S. Angrist and Lavy (1999) point out that the magnitude found by these authors is nevertheless also relatively close to theirs.

6 Conclusions

Development practitioners are increasingly emphasizing education as a key component of strategies to achieve growth and equity. Further, they are keen to raise schooling quality, both because several middle income countries have already achieved a number of enrollment goals, and because this may be an effective way to help low income countries achieve theirs.

Hanushek (1995) makes the point, however, that there may be little economists can reliably tell them on how to allocate scarce resources among educational inputs. This shortcoming is particularly severe if one considers that, ideally, cost-benefit studies require knowledge of the magnitude of certain effects, and not merely their direction.

This paper makes the point that the endogeneity issues that cause this problem may be particularly severe in developing countries. Specifically, factors like highly unequal income distributions, ethnic heterogeneity, significant urban/rural welfare differences, and the presence of "open" public school systems, could lead to a great deal of sorting among schools, complicating analyses on any number of inputs.

In the case analyzed here, for instance, the evidence is of positive enrollment/class size and enrollment/SES correlations, which jointly imply that higher SES and higher achieving students may often be not only in larger schools, but also in larger classes, which by itself can produce class size coefficients of the "wrong" sign. This problem underlines the importance of identifying exogenous variation in input provision in order to generate reliable results.

This paper has attempted to do this in the case of class size, building two empirical strategies that are based on the institutional characteristics of educational provision in the rural area of Bolivia. When these are implemented, the data reveal negative and significant class size effects, not so different in magnitude from those found in recent experimental or quasi-experimental studies. These significant results might serve as a starting point for further work on this topic in the Bolivian setting, which would ideally include experimental designs.

References

Akerhielm, Karen, "Does Class Size Matter?," *Economics of Education Review*, 1995, *XIV*, 229–241.

Albó, Javier, *Bolivia Plurilingue: Guia para Planificadores y Educadores*, La Paz: Unicef-Cipca, 1995.

Angrist, Joshua D. and Victor Lavy, "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics*, 1999, pp. 533–575.

Behrman, Jere and Nancy Birdsall, "The Quality of Schooling: Quantity Alone is Misleading," *American Economic Review*, 1983, *73* (5), 928–946.

Boozer, Michael and Cecilia Rouse, "Intraschool variation in class size: patterns and implications," National Bureau of Economic Research Working Paper No. 5144 1995.

Campbell, D.T., "Reforms as Experiments," *American Psychologist*, 1969, *24*, 409–429.

Card, David and Alan Krueger, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, 1992, *100* (1), 1–40.

Case, Anne and Angus Deaton, "School Quality and Educational Outcomes in South Africa," Research Program in Development Studies, Princeton University, Mimeo 1998.

___ **and M. Yogo**, "Does School Quality Matter? Returns to Education and the Characteristics of Schools in South Africa," Research Program in Development Studies, Princeton University, Mimeo 1999.

Coleman, James and et Al., "Equality of Educational Opportunity," Washington, Goverment Printing Office 1966.

Finn, Jeremy D. and Charles M. Achilles, "Answers and Questions about Class Size: a Statewide Experiment," *American Educational Research Journal*, 1990, *Fall*, 557–577.

___ **and** ___ , "Tennesse's Class Size Study: Findings, Implications, Misconceptions," *Educational Evaluation and Policy Analysis*, 1999, *21* (2), 97–109.

Fuller, Bruce, "Raising School Quality in Developing Countries, What Investments Boost Learning," World Bank Discussion Papers 1986.

___ **and Prema Clarke**, "Raising School Effects While Ignoring Culture? Local Conditions and the Influence of Classroom Tools, Rules, and Pedagogy," *Review of Educational Research*, 1994, *64* (1), 119–157.

Hanushek, Erik A., "The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature*, 1986, *24*, 1141–1177.

___ , "Interpreting Recent Research on Schooling in Developing Countries," *World Bank Research Observer*, 1995, *10* (2), 227–246.

___ , "The Evidence on Class Size," W. Allen Wallis Institute of Political Economy Occasional Working Paper 98-1. University of Rochester. 1998.

___ , "Some Findings from an Independent Investigation of the Tennessee STAR Experiment and from other Investigations of Class Size Effects," *Educational Evaluation and Policy Analysis*, 1999, *21* (2), 143–163.

Hedges, Larry V., Richard D. Laine, and Rob Greenwald, "Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes," *Educational Researcher*, 1994, *23* (3), 5–14.

Hoxby, Caroline Minter, "Does Competition among Public Schools Benefit Students and Taxpayers? Evidence from Natural Variation in School Districting," mimeo 1997.

___ , "The effects of class size and composition on student achievement: new evidence from natural population variation," 1999.

Kremer, Michael, "Research on Schooling: What We Know and What We Don't. A Comment on Hanushek," *World Bank Research Observer*, 1995, *10*, 247–254.

Krueger, Alan B., "Experimental Estimates of Education Production Functions," *Quartely Journal of Economics*, 1999, pp. 497–531.

Lavy, Victor, "Endogenous School Resources and Cognitive Achievement in Primary Schools in Israel," Hebrew University, Falk Institute Discussion Paper No. 95.03. 1995.

___ , "Using Dual Natural Quasi-Experimental Designs to Evaluate the Effect of School Hours and Class Size on Student Achievement," Mimeo 1999.

Ministerio de Desarrollo Humano, "Rendimientos escolares de 3ro. y 6to. de primaria en Lenguaje y Matemática y factores asociados," Sistema de Medición de Calidad, Ministerio de Educación Cultura y Deportes. 1997.

Mizala, Alejandra, Pilar Romaguera, and Teresa Reinaga, "Factores que Inciden en el Rendimiento Escolar en Bolivia," Universidad de Chile, Mimeo. 1999.

Moesteller, Frederick, "The Tennessee Study of Class Size in the Early School Grades," *The Future of Children: Critical Issues for Children and Youths*, 1995, *V*, 113–127.

Urquiola, Miguel S., "Demand Matters: School District Concentration, Composition, and Educational Expenditure," The World Bank 2000.

Table 1:

Descriptive statistics by type of school

	Total	Urban private	Urban public	Rural public
Test scores:				
Sample size	10,018	1,139	5,285	3,594
Language score: mean	49.8	58.2	49.8	47.2
Language score: std. dev.	9.3	8.2	8.4	9.3
Math score: mean	53.2	60.8	53.2	50.9
Math score: std. dev.	9.2	6.9	8.4	9.7
Class characteristics:				
Sample size	608	46	252	310
3rd grade (school) enrollment: mean	63.3	77.6	91.7	38.2
3rd grade (school) enrollment: std. dev.	44.1	59.6	39.0	27.4
3rd grade class size: mean	29.9	37.4	35.5	23.3
3rd grade class size: std. dev.	9.9	9.8	6.4	8.4
No. of 3rd grade classes in school: mean	1.6	1.4	2.2	1.3
% of classes with blackboard[1]	95.2	100.0	94.5	94.3
% of students with textbooks[2]	36.2	99.1	40.3	23.4
% of classrooms with libraries[3]	18.5	39.5	9.9	22.4
Teachers' characteristics:				
Sample size	608	46	252	310
Mean age	40.0	45.1	41.7	38.1
% with Spanish as first language	80.0	89.2	85.5	74.5
% with degree	79.3	97.6	82.5	74.1
% in categories 1-2 (top)[4]	21.1	46.4	25.5	14.0
% in categories 3-4[4]	24.3	33.2	27.8	20.4
% in categories 5-6[4]	19.7	16.0	27.4	14.6
% in category 7 or without category[4]	34.9	4.4	19.3	51.0
Students' characteristics:				
Sample size	10,018	1,139	5,285	3,594
Mean age	9.2	8.7	9.1	9.5
% with Spanish as first language	75.0	93.3	81.2	60.2
% that works at least occasionally	67.9	40.2	62.9	84.1
% whose mother finished at least grade school	87.3	99.5	90.6	78.6
% whose mother finished at least middle school	47.1	91.2	50.9	27.5
% whose mother finished at least high school	30.6	82.5	31.4	12.9
% whose mother has at least some higher ed.	14.0	53.9	11.7	4.8
% whose mother has at least college degree	3.8	20.1	2.4	0.8
% whose hhld. has sewage system	30.7	74.1	35.8	9.6
% whose hhld. has a phone line	17.5	71.7	15.9	2.7

Note: The sample size refers to the total number of observations in each category. Not all variables had valid data for that many observations.

[1] The teacher responds he/she always has these inputs.

[2] The teacher responds that more than half of students have textbooks.

[3] The teacher indicates his/her classroom has a set of reference textbooks.

[4] Categories correspond to teachers' position in the pay scale.

Table 2:

Dependent variable: language score
Sample: 3rd grade

	Full sample		Urban private	Urban public	Rural public
	(1)	(2)	(3)	(4)	(5)
Panel A					
Constant	46.8***	61.6***	54.5***	55.6***	49.9***
	(1.2)	(1.6)	(2.8)	(2.2)	(1.4)
Class size	0.093**	-0.086**	0.093	-0.156**	-0.102*
	(0.037)	(0.036)	(0.065)	(0.060)	(0.057)
	[0.09]	[-0.08]	[0.11]	[-0.12]	[-0.09]
Urban Public		-8.6***			
		(0.9)			
Rural Public		-12.1***			
		(1.1)			
R^2	0.008	0.125	0.012	0.013	0.009
No. of students	9,993	9,993	1,139	5,260	3,594
No. of classes	589	589	46	245	298
Panel B					
Constant	49.0***	59.1***	57.1***	51.6***	48.2***
	(0.7)	(1.0)	(1.4)	(1.0)	(0.7)
Enrollment	0.011	-0.011*	0.013	-0.018*	-0.024*
	(0.007)	(0.007)	(0.012)	(0.009)	(0.013)
	[0.05]	[-0.05]	[0.09]	[-0.08]	[-0.07]
Urban Public		-8.2***			
		(0.9)			
Rural Public		-11.4***			
		(1.0)			
R^2	0.003	0.122	0.008	0.007	0.005
No. of students	9,993	9,993	1,139	5,260	3,594
No. of classes	589	589	46	245	298

*, **, *** - significant at the 10, 5 and 1 percent level, respectively.
1) Huber-White standard errors are in parentheses.
2) Brackets contain the proportion of a standard deviation change in the dependent variable brought about by increasing the independent variable by one standard deviation.
3) Standard errors are adjusted for clustering at the class level. See Moulton (1986).

Table 3:

Dependent variable: language score
Sample: 3rd grade

	Full sample		Urban private		Urban public		Rural public	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	46.8***	49.3***	54.5***	46.9***	55.6***	53.7***	49.9***	51.3***
	(1.2)	(1.1)	(2.8)	(3.5)	(2.2)	(1.9)	(1.4)	(2.0)
Class size	0.093**	-0.005	0.093	0.052	-0.156**	-0.099*	0.102*	-0.088
	(0.037)	(0.033)	(0.065)	(0.060)	(0.060)	(0.052)	(0.057)	(0.083)
	[0.09]	[-0.01]	[0.11]	[0.06]	[-0.12]	[-0.07]	[-0.09]	[-0.08]
Enrollment		-0.007		0.005		-0.012		-0.016
		(0.007)		(0.009)		(0.009)		(0.021)
Student speaks Spanish only		1.80***		2.20		1.59***		1.28*
		(0.43)		(1.31)		(0.52)		(0.75)
Student works at least part time		-2.60***		-1.10		-2.62***		-1.78**
		(0.31)		(0.77)		(0.35)		(0.76)
Mother completed grade school		-0.61		0.24		-0.48		-1.23**
		(0.39)		(2.41)		(0.41)		(0.60)
Mother completed middle school		-0.22		-0.11		0.22		-0.70
		(0.45)		(2.41)		(0.48)		(0.76)
Mother completed high school		1.50***		4.32		0.96*		1.61*
		(0.47)		(2.74)		(0.53)		(0.85)
Mother has some higher ed.		3.41***		5.72**		2.67***		1.53
		(0.52)		(2.73)		(0.64)		(1.03)
Mother is a college graduate		4.84***		6.67**		2.81***		0.52
		(0.66)		(2.82)		(1.06)		(1.68)
Hhld. has access to sewage sys.		1.66***		1.53		0.99**		1.68*
		(0.35)		(0.99)		(0.41)		(0.97)
Hhld. has a phone connection		3.72***		2.19***		2.24***		1.03
		(0.44)		(0.80)		(0.42)		(1.18)
R^2	0.008	0.158	0.012	0.172	0.013	0.093	0.009	0.037
No. of students	9,993	9,993	1,139	1,139	5,260	5,260	3,594	3,594
No. of classes	589	589	46	46	245	245	298	298

*, **, *** - significant at the 10, 5 and 1 percent level, respectively.
1) Huber-White standard errors are in parentheses.
2) Brackets contain the proportion of a standard deviation change in the dependent variable brought about by increasing the independent variable by one standard deviation.
3) Standard errors are adjusted for clustering at the class level. See Moulton (1986).

Table 4:

Reduced form regressions for *rural public schools*
Sample: 3rd grade

	Class size		Language score		Math score	
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	4.85***	6.79***	51.2***	53.1***	55.4***	56.1***
	(1.27)	(1.58)	(1.6)	(1.7)	(1.6)	(1.8)
Predicted class size	0.84***	0.65***	-0.155***	-0.190**	-0.177**	-0.146*
	(0.04)	(0.07)	(0.059)	(0.079)	(0.059)	(0.083)
			[-0.12]	[-0.15]	[-0.13]	[-0.11]
Enrollment		0.072***		0.000		-0.024
		(0.016)		(0.018)		(0.019)
Student speaks Spanish only		0.20		1.20		1.42**
		(0.56)		(0.73)		(0.68)
Student works at least part time		-0.25		-1.79**		-1.49**
		(0.28)		(0.77)		(0.70)
Mother completed grade school		0.00		-1.18*		-0.91
		(0.38)		(0.61)		(0.64)
Mother completed middle school		0.14		-0.63		-0.26
		(0.54)		(0.78)		(0.77)
Mother completed high school		-0.02		1.76**		0.96
		(0.65)		(0.88)		(0.88)
Mother has some higher ed.		-1.02		1.80*		2.69***
		(0.69)		(1.04)		(1.02)
Mother is a college graduate		-1.68**		0.95		2.65*
		(0.81)		(1.69)		(1.60)
Hhld. has access to sewage system		-0.53		1.83*		1.76**
		(0.36)		(0.98)		(0.83)
Hhld. has a phone connection		-0.04		1.09		2.69***
		(0.50)		(1.19)		(1.02)
R^2	0.520	0.552	0.014	0.044	0.017	0.047
No. of students	3,594	3,594	3,594	3,594	3,594	3,594
No. of classes	298	298	298	298	298	298

*, **, *** - significant at the 10, 5 and 1 percent level, respectively.
1) Huber-White standard errors are in parentheses.
2) Brackets contain the proportion of a standard deviation change in the dependent variable brought about by increasing the independent variable by one standard deviation.
3) Standard errors are adjusted for clustering at the class level. See Moulton (1986).

Table 5:

Instrumental variables regressions for *rural* public schools
Sample: 3rd grade

	Language score				Math score			
	OLS		IV		OLS		IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	49.9***	51.3***	52.1***	55.1***	54.1***	54.7***	56.5***	57.7***
	(1.4)	(2.0)	(1.9)	(2.4)	(1.4)	(1.7)	(1.9)	(2.6)
Class size	-0.102*	-0.088	-0.184***	-0.292**	-0.121**	-0.062	-0.210***	-0.224*
	(0.057)	(0.083)	(0.070)	(0.122)	(0.053)	(0.073)	(0.070)	(0.128)
	[-0.09]	[-0.08]	[-0.17]	[-0.26]	[-0.11]	[-0.05]	[-0.18]	[-0.20]
Enrollment		-0.016		0.022		-0.038		-0.008
		(0.021)		(0.025)		(0.019)		(0.027)
Student speaks Spanish only		1.28*		1.25*		1.49**		1.47**
		(0.75)		(0.74)		(0.69)		(0.69)
Student works at least part time		-1.78**		-1.86**		-1.48**		-1.55**
		(0.76)		(0.77)		(0.69)		(0.71)
Mother completed grade school		-1.23**		-1.18*		-0.95		-0.91
		(0.60)		(0.61)		(0.63)		(0.64)
Mother completed middle school		-0.70		-0.59		-0.32		-0.23
		(0.76)		(0.78)		(0.77)		(0.77)
Mother completed high school		1.61*		1.75**		0.84		0.96
		(0.85)		(0.88)		(0.87)		(0.88)
Mother has some higher ed.		1.53		1.50		2.49**		2.46**
		(1.03)		(1.05)		(1.00)		(1.01)
Mother is a college graduate		0.52		0.46		2.32		2.27
		(1.68)		(1.71)		(1.62)		(1.64)
Hhld. has access to sewage sys.		1.68*		1.68*		1.64**		1.64**
		(0.97)		(0.98)		(0.83)		(0.83)
Hhld. has a phone connection		1.03		1.08		2.65***		2.68***
		(1.18)		(1.19)		(1.01)		(1.02)
R^2	0.009	0.037	0.014	0.044	0.011	0.043	0.017	0.047
No. of students	3,594	3,594	3,594	3,594	3,594	3,594	3,594	3,594
No. of classes	298	298	298	298	298	298	298	298

*, **, *** - significant at the 10, 5 and 1 percent level, respectively.
1) Huber-White standard errors are in parentheses.
2) Brackets contain the proportion of a standard deviation change in the dependent variable brought about by increasing the independent variable by one standard deviation.
3) Standard errors are adjusted for clustering at the class level. See Moulton (1986).

Table 6:

Regressions for *all rural schools* and *rural schools with enrollments of 30 or less*
Sample: 3rd grade

	Language score				Math score			
	All Rural Schools		Schools with enrollments below 30		All Rural Schools		Schools with enrollments below 30	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	49.9***	51.3***	51.9***	53.8***	54.1***	54.7***	55.8***	56.7***
	(1.4)	(2.0)	(2.1)	(2.4)	(1.4)	(1.7)	(2.1)	(2.6)
Class size	-0.102*	-0.088	-0.218**	-0.225**	-0.121**	-0.062	-0.189**	-0.194**
	(0.057)	(0.083)	(0.102)	(0.098)	(0.053)	(0.073)	(0.092)	(0.099)
	[-0.09]	[-0.08]	[-0.13]	[-0.14]	[-0.11]	[-0.05]	[-0.12]	[-0.12]
Enrollment		-0.016				-0.038		
		(0.021)				(0.019)		
Student speaks Spanish only		1.28*		1.53		1.49**		1.67
		(0.75)		(0.96)		(0.69)		(1.07)
Student works at least part time		-1.78**		-2.40**		-1.48**		-1.93
		(0.76)		(1.22)		(0.69)		(1.23)
Mother completed grade school		-1.23**		-1.00		-0.95		-0.49
		(0.60)		(0.84)		(0.63)		(0.94)
Mother completed middle school		-0.70		-1.63		-0.32		0.07
		(0.76)		(1.11)		(0.77)		(1.21)
Mother completed high school		1.61*		0.88		0.84		0.12
		(0.85)		(1.33)		(0.87)		(1.54)
Mother has some higher ed.		1.53		1.07		2.49**		2.83*
		(1.03)		(1.61)		(1.00)		(1.51)
Mother is a college graduate		0.52		2.34		2.32		3.23
		(1.68)		(2.30)		(1.62)		(3.86)
Hhld. has access to sewage sys.		1.68*		0.71		1.64**		1.90
		(0.97)		(2.30)		(0.83)		(2.42)
Hhld. has a phone connection		1.03		0.50		2.65***		1.27
		(1.19)		(1.82)		(1.01)		(1.37)
R^2	0.009	0.037	0.018	0.041	0.011	0.043	0.012	0.032
No. of students	3,594	3,594	1,555	1,555	3,594	3,594	1,555	1,555
No. of classes	298	298	159	159	298	298	159	159

*, **, *** - significant at the 10, 5 and 1 percent level, respectively.

1) Huber-White standard errors are in parentheses.

2) Brackets contain the proportion of a standard deviation change in the dependent variable brought about by increasing the independent variable by one standard deviation.

3) Standard errors are adjusted for clustering at the class level. See Moulton (1986).

Table 7:

Rural public schools with enrollments under 31 *and* those "close" to discontinuities
Sample: 3rd grade

	Language score				Math score			
	OLS		IV		OLS		IV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	49.9***	51.1***	53.2***	56.9***	54.6***	55.3***	57.1***	59.3***
	(2.2)	(3.0)	(2.2)	(2.7)	(1.9)	(2.3)	(2.2)	(2.9)
Class size	-0.105	-0.134	-0.250***	-0.545**	-0.128	-0.132	-0.233**	-0.417**
	(0.104)	(0.136)	(0.088)	(0.218)	(0.080)	(0.109)	(0.094)	(0.179)
	[-0.09]	[-0.11]	[-0.21]	[-0.45]	[-0.10]	[-0.10]	[-0.18]	[-0.33]
Enrollment		0.025		0.170		0.002		0.102
		(0.060)		(0.110)		(0.058)		(0.082)
Student speaks Spanish only		2.00		1.58		1.57*		1.28
		(1.03)		(1.00)		(0.92)		(0.96)
Student works at least part time		-1.86*		-2.23*		-1.19		-1.46
		(1.05)		(1.14)		(1.05)		(1.14)
Mother completed grade school		-1.61**		-1.21		-1.21		-0.93
		(0.69)		(0.79)		(0.77)		(0.79)
Mother completed middle school		-1.41		-0.86		-0.86		-0.48
		(1.01)		(1.31)		(1.05)		(1.09)
Mother completed high school		1.24		1.98		0.72		1.23
		(1.13)		(1.37)		(1.29)		(1.36)
Mother has some higher ed.		1.27		1.35		3.03**		3.09**
		(1.43)		(1.58)		(1.27)		(1.38)
Mother is a college graduate		-1.13		0.92		2.18		3.61
		(2.79)		(3.02)		(3.57)		(3.52)
Hhld. has access to sewage system		2.44*		2.71*		1.87		2.06
		(1.32)		(1.38)		(1.40)		(1.39)
Hhld. has a phone connection		0.75		0.83		2.13		2.20
		(1.73)		(1.94)		(1.34)		(1.59)
R^2	0.008	0.045	0.026	0.080	0.010	0.036	0.021	0.052
N	2,252	2,252	2,252	2,252	2,252	2,252	2,252	2,252
No. of classes	209	209	209	209	209	209	209	209

*, **, *** - significant at the 10, 5 and 1 percent level, respectively.
1) Huber-White standard errors are in parentheses.
2) Brackets contain the proportion of a standard deviation change in the dependent variable brought about by increasing the independent variable by one standard deviation.
3) Standard errors are adjusted for clustering at the class level. See Moulton (1986).

Table 8:

Dependent variable: grade-specific language score
Within town variation
Class-level observations

	(1)	(2)	(3)	(4)	(5)	(6)
Constant	51.9***	51.8***	49.3***	44.4***	44.1***	50.4***
	(1.9)	(2.1)	(1.5)	(4.8)	(4.9)	(6.7)
Class size	-0.224**	-0.223**	-0.183**	-0.119	-0.113	-0.181**
	(0.091)	(0.092)	(0.079)	(0.088)	(0.089)	(0.078)
	[-0.18]	[-0.18]	[-0.14]	[-0.09]	[-0.09]	[-0.14]
6^{th} grade		0.154	0.337		0.563	0.345
		(1.169)	(0.533)		(1.120)	(1.638)
Student speaks Spanish only				6.57***	6.59***	5.17
				(1.47)	(1.47)	(4.20)
Student works at least part time				-6.20*	-6.35*	-0.75
				(3.58)	(3.60)	(5.40)
Mother completed grade school				9.26***	9.41***	-3.21
				(3.43)	(3.44)	(4.48)
Mother completed middle school				9.91**	9.78**	-6.99
				(4.28)	(4.30)	(6.43)
Mother completed high school				4.97	5.07	-3.14
				(6.80)	(6.82)	(10.92)
Mother has some higher ed.				13.29	13.68	2.13
				(11.70)	(11.75)	(14.31)
Mother is a college graduate				25.82	26.60	27.94
				(44.16)	(44.28)	(48.90)
Hhld. has access to sewage system				-7.67	-7.67	-2.28
				(11.56)	(11.59)	(17.24)
Hhld. has a phone connection				0.55	0.67	-2.29
				(5.53)	(5.55)	(16.07)
Town dummies	No	No	**Yes**	No	No	**Yes**
R^2	0.031	0.032	0.903	0.227	0.228	0.914
N	190	190	190	190	190	190

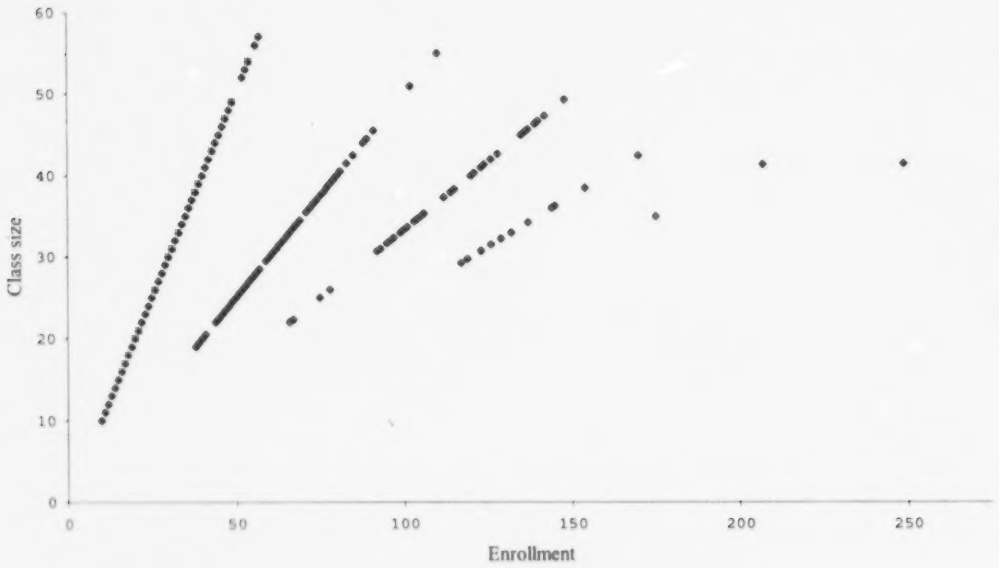
*, **, *** - significant at the 10, 5 and 1 percent level, respectively.
1) Huber-White standard errors are in parentheses.
2) Brackets contain the proportion of a standard deviation change in the dependent variable brought about by increasing the independent variable by one standard deviation.
3) Standard errors are adjusted for clustering at the class level. See Moulton (1986).

Table 9:

Summary of effect sizes and comparisons

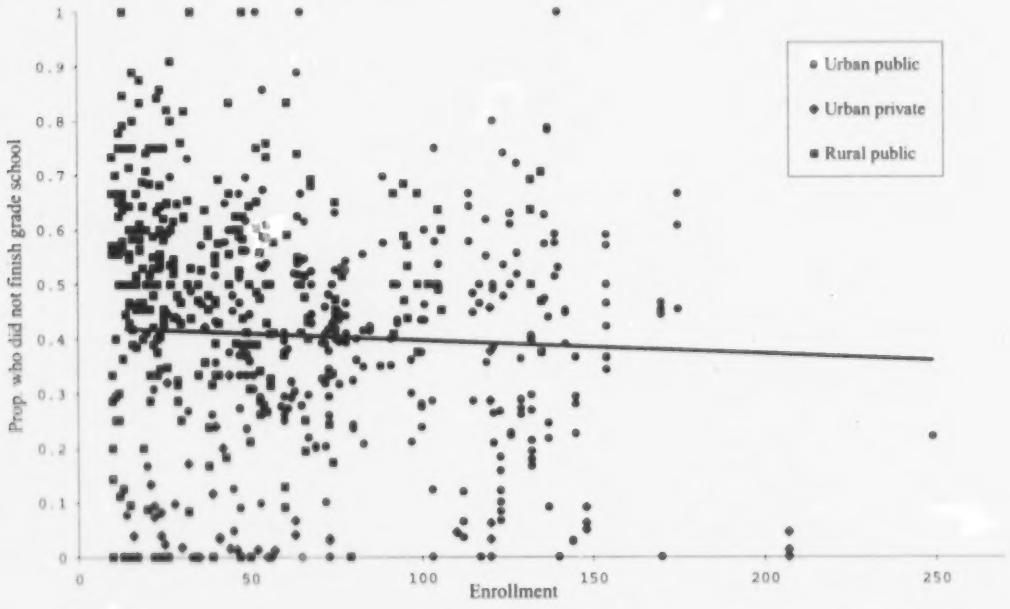
Study/estimate/effect	Relative to the student distribution	Relative to the class mean distribution
Finn and Achilles (1990)	0.13σ-0.27σ	0.32σ-0.66σ
Angrist and Lavy (1999)	0.10σ-0.20σ	0.36σ
This study: IV estimates	0.17σ-0.26σ	0.34σ-0.55σ
This study: "small" school estimates	0.18σ-0.23σ	0.19σ-0.37σ
This study: "mixed" strategy	0.18σ-0.45σ	0.20σ-0.39σ

Figure 1: Enrollment and class size



Note: 1) Plot of class size and enrollment, with the former computed according to formula (1).
2) The figure uses class-level observations for the 3rd. grade.

Figure 2: Enrollment and mothers' schooling



Note: The proportion of mothers with fewer than five years of schooling is on the Y axis.
The figure uses class-level observations for the 3rd. grade.

Figure 3
Enrollment: relative freq. distribution by type of school

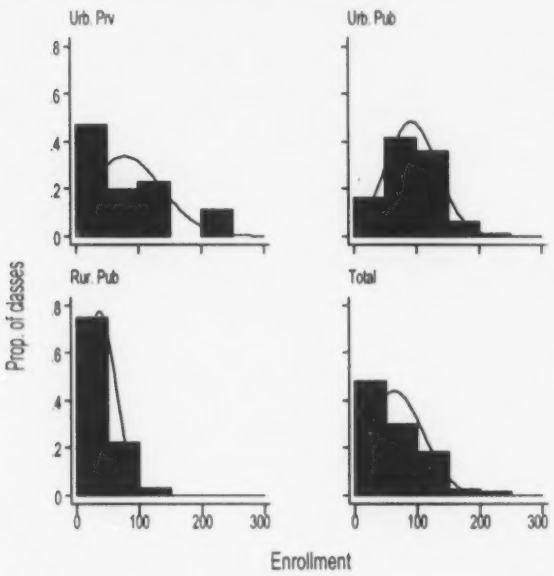


Figure 4:
Prop. of children who speak only Spanish: rel. freq. dist.

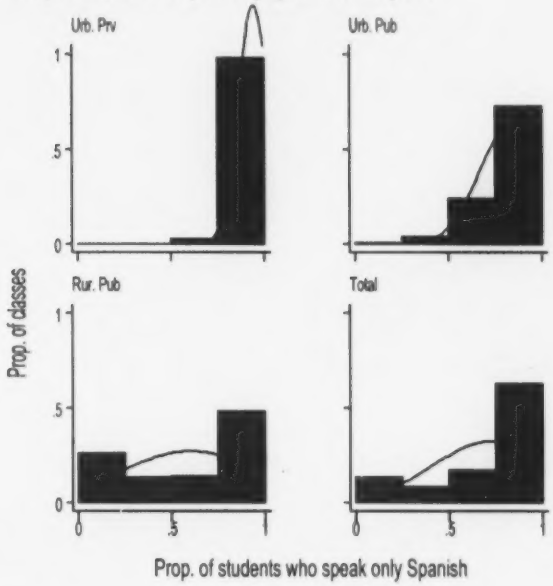


Figure 5:
Prop. of students with texts: rel. freq. dist. by type of school

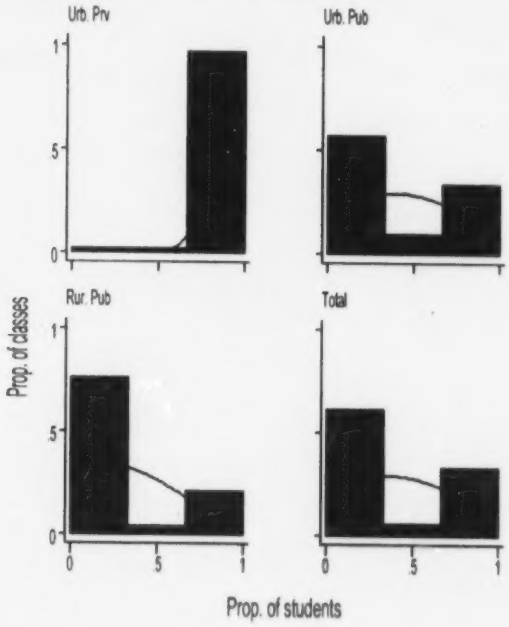


Figure 6:
Prop. of teachers at the top of the pay scale: rel. freq. dist.

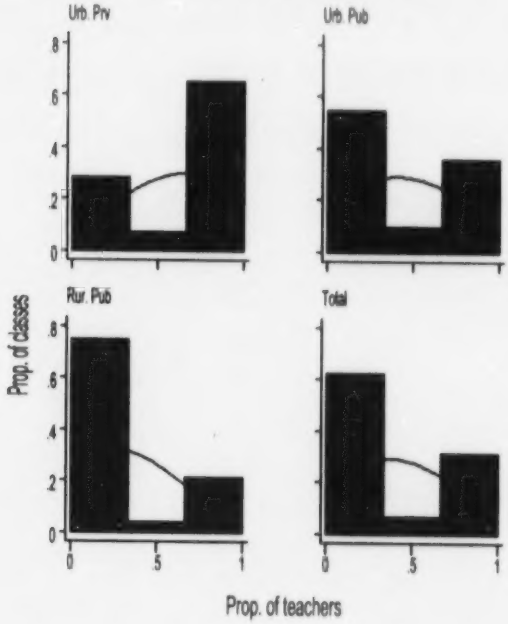
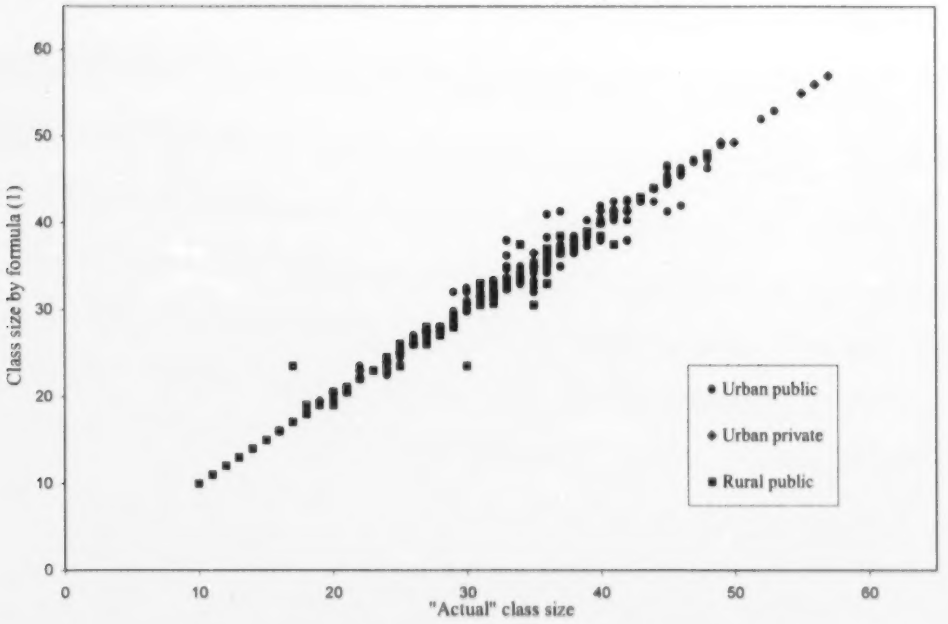
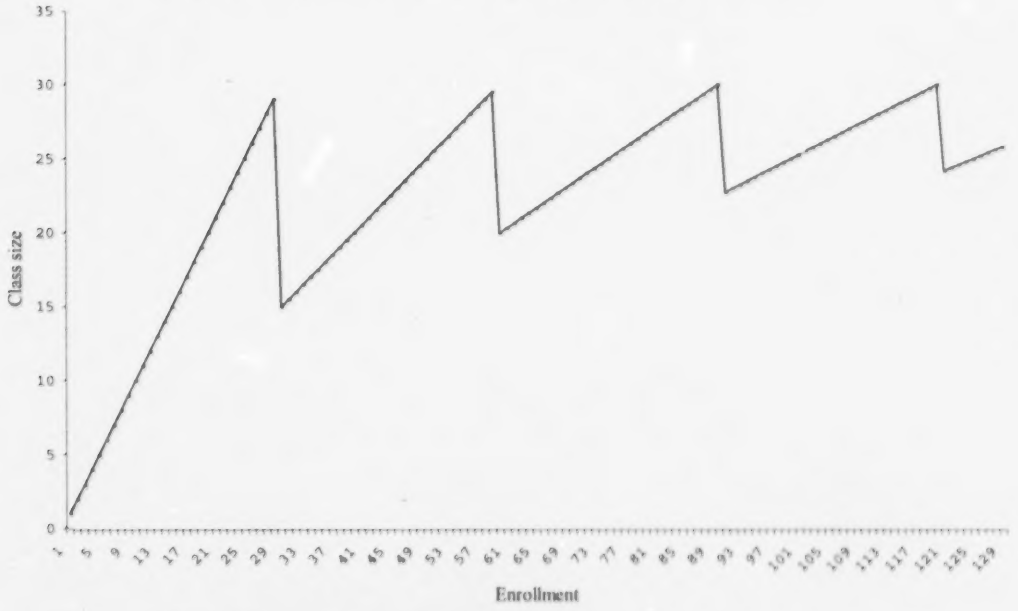


Figure 7: Measurement error in class size



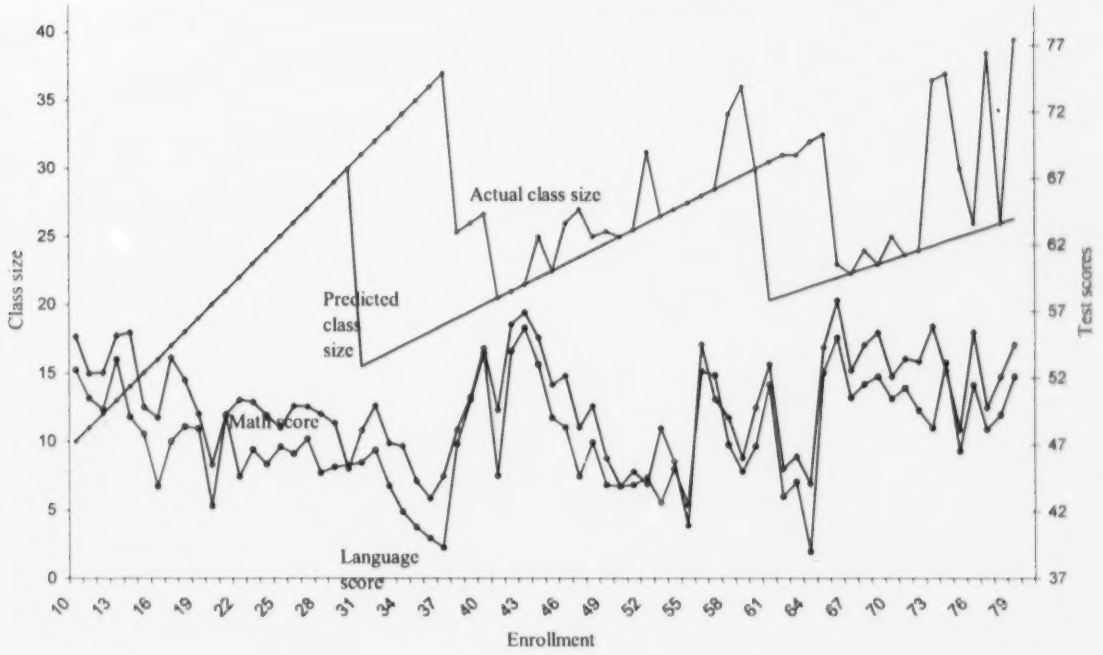
Note: 1) The X axis features class size as found by the testing system (based on the teachers declaration).
The Y axis displays the ratio of enrollment and the no. of classes.
2) The figure uses class-level observations for the 3rd. grade.

Figure 8: Enrollment and class size under the teacher allocation rule



Note: 1) Plot of predicted class size given enrollment, according to formula (5).
2) The figure uses class-level observations for the 3rd. grade.

Figure 9: Enrollment, predicted and actual class size; math and language scores



Note: 1) The figure plots **averages** (by enrollment level) of: actual and predicted class sizes; and of math and language scores.
2) The figure uses class-level observations for the 3rd. grade.